Evidence for a duplicated mitochondrial region in Audubon's shearwater based on MinION sequencing

Torres Lucas[1,2*], Welch Andreanna J.[3], Zanchetta Catherine[4], Chesser R. Terry [5], Manno Maxime[4], Donnadieu Cécile[4], Bretagnolle Vincent[1], Pante Eric[2]

Affiliations

1 Centre d'Etudes Biologiques de Chizé, UMR 7372, CNRS & Université de La Rochelle, Villiers en Bois, France

2 Littoral, Environnement et Sociétés, UMR 7266 CNRS, Université de La Rochelle, La Rochelle, France

3 Department of Biosciences, Durham University, South Road, Durham, DH1 3LE, UK

4 US1426 Get-PlaGe, Centre INRA de Toulouse Midi-Pyrénées, Castanet-Tolosan, France

5 USGS Patuxent Wildlife Research Center, National Museum of Natural History, Smithsonian Institution, Washington, DC 20013

*Corresponding author

Correspondence details and Biographical notes:

Lucas Torres:

I am a PhD student studying phylogeography of several petrel complexes and molecular evolution in the Procellariiformes.

(+33) 05 49 09 35 51

lucas.torres@cebc.cnrs.fr

Centre d'Etudes Biologiques de Chizé

CNRS UMR 7372 - Université de La Rochelle

405 Route de La Canauderie

79360 Villiers-en-Bois

Andreanna J. Welch:

My research interests center broadly on gaining a better understanding of the amazing biodiversity that we see around us today, and how this biodiversity has changed through time. I work on multiple scales at the intersection between ecology, evolution, cellular biology, physiology, and conservation.

+44 (0)191-334-1258

a.j.welch@durham.ac.uk

Department of Biosciences, Durham University, South Road, Durham, DH1 3LE, UK

Catherine Zanchetta:

I am a molecular biologist at the Sequencing Core Facility of Toulouse (GeT-PlaGe - INRA). The mission of the core facility is to provide innovating technologies for genome and transcriptome analysis to the scientific community. I work on Nanopore and Illumina technologies. Prior to working at GeT-PlaGe, I completed a Master's degree in Molecular diagnostic at the University Paul Sabatier in Toulouse, France.

+33 (0)5 61 28 51 76

catherine.zanchetta@gmail.com

GeT-PlaGe
Campus INRA

24 chemin de borde rouge - Auzeville
CS 52627
31326 CASTANET-TOLOSAN Cedex

R. Terry Chesser:

My research focuses on avian systematics, evolution, and phylogeography.

(202) 633-4886

chessert@si.edu

USGS Patuxent Wildlife Research Center

National Museum of Natural History

Smithsonian Institution

PO Box 37012, MRC 111

Washington, DC 20013-7012

Maxime Manno:

Currently, I am working at GeT-PlaGe (Genomic core facility at INRA-Toulouse) as engineer in Bioinformatics. I am in charge of the informatics and bioinformatics part of the generation of ONT data and the quality control. I contributed to the installation of the ONT technology (MinION and GridION) on the core facility, the test of software (as Albacore, porechop) and the development of the IT solution.

+33 (0)5 61 28 57 47

maxime.manno@inra.fr

GeT-PlaGe
Campus INRA
24 chemin de borde rouge - Auzeville
CS 52627
31326 CASTANET-TOLOSAN Cedex

Cécile Donnadieu:

I currently work at the Institut national de la recherche agronomique (Toulouse). I do research in Molecular Biology, Bioinformatics and Genetics.

+33 (0)5 61 28 57 54

cecile.donnadieu@inra.fr

GeT-PlaGe
Campus INRA
24 chemin de borde rouge - Auzeville
CS 52627
31326 CASTANET-TOLOSAN Cedex

Vincent Bretagnolle:

My research is part of the Agro-ecology of territories, and focus in particular on the analysis of ecological services rendered by biodiversity "both economic and socio-cultural" to promote them. My work is carried out mainly on the Plaine & Val de Sèvre workshop area (department of Deux Sèvres) that I created in 1994. I have also been studying Procellariiformes seabirds for nearly 30 years.

(+33) 05 49 09 78 17

breta@cebc.cnrs.fr

Centre d'Etudes Biologiques de Chizé

CNRS UMR 7372 - Université de La Rochelle

405 Route de La Canauderie

79360 Villiers-en-Bois

Eric Pante:

I am broadly interested in the molecular ecology and evolution of marine organisms.

+33 (0)5 46 50 65 43

eric.pante@univ-lr.fr

Institut du Littoral et de l'Environnement

2, rue Olympe de Gouges

17 000 La Rochelle

Abstract

Mitochondrial genetic markers have been extensively used to study the phylogenetics and phylogeography of many birds, including seabirds of the order Procellariiformes. Evidence suggests that part of the mitochondrial genome of Procellariiformes, especially albatrosses, is duplicated, but no DNA fragment covering the entire duplication has been sequenced. We sequenced the complete mitochondrial genome of a non-albatross species of Procellariiformes, *Puffinus lherminieri* (Audubon's shearwater) using the long-read MinION (ONT) technology. Two mito-genomes were assembled from the same individual, differing by 52 SNPs and in length. The shorter was 19kb-long while the longer was 21 kb, due to the presence of two identical copies of *nad6*, three tRNA, and two dissimilar copies of the control region. Contrary to albatrosses, *cob* was not duplicated. We further detected a complex repeated region of undetermined length between the control region and 12S. Long-read sequencing suggests heteroplasmy and a novel arrangement within the duplicated region, indicating a complex evolution of the mitogenome in Procellariiformes.

Keywords: Heteroplasmy; Control Region; Long-Range PCR; Tandem repeats; Cytochrome-b; cytb

Table of Contents

Introduction

Many studies have been conducted on the evolutionary biology and phylogeography of Procellariiformes, a group of seabirds (albatrosses and petrels) that, like other seabirds, have high dispersal abilities and tend to be distributed over vast areas (Harrison 2000). Most such studies were based on mitochondrial markers, principally *cox1*, *cob* and the Control Region (CR). MtDNA has been used for phylogenetic (e.g., Austin et al. 2004; Jesus et al. 2009; Welch et al. 2014), phylogeographic (e.g., Cagnon et al. 2004; Smith et al. 2007; Kerr & Dove 2013) and population genetics studies (e.g., Burg & Croxall 2001; Alderman et al. 2005; Ramírez et al. 2013) but use of mtDNA for such purposes relies on the assumption that the markers are present in a single copy (e.g. Brown 1985; Avise et al. 1987; Moritz et al. 1987; Boore 1999). However, it is known that heteroplasmy (e.g. Berg et al. 1995, Mundy et al. 1996, Moum & Bakke 2001, Gandolfi et al. 2017), mitochondrial pseudogenes or NUMTS (Sorenson & Quinn 1998), and recombination (e.g. Tsaousis et al. 2005, Sammler et al. 2011) occur in birds. For instance, Abbott et al. (2005) found evidence of a duplicated region in the mitochondrial genome of albatrosses, resulting in two copies of *nad6*, CR, and two fragments of *cob*. Two divergent copies of the mitochondrial control region have since been indirectly suggested in eight additional species of Procellariiformes (Smith et al. 2007; Lawrence et al. 2008, 2014; Burg et al. 2014; Rains et al. 2011; Burg et al. 2014), covering three of the four Procellariiformes families (Gangloff et al. 2013; Welch et al. 2014; Prum et al. 2015). The partial duplication of the mitochondrial genome is therefore apparently widespread within the Procellariiformes. However, apart from the study of Abbott et al. (2005), only three studies, all of albatrosses, have sequenced the complete duplicated region. Gibb et al. (2007) revised a previously-published genome of *Thalassarche melanophris* (AY158677, Slack et al. 2006), Eda et al. (2010) used primer-walking to sequence the duplicated region of three species from the genus *Phoebastria*, and Lounsberry *et*

al. (2015) sequenced the complete mitochondrial genome of these same three *Phoebastria* species using Illumina and Sanger sequencing. Nearly-complete mitochondrial genomes of *Diomedea chrysostoma*, *Procellaria cinerea*, and *Pterodroma brevirostris*, which apparently lack the duplication, were also sequenced (Slack et al. 2006, Watanabe et al. 2006) (Tab. 1). Here we sequenced the complete mitochondrial genome of *Puffinus lherminieri*, providing the first complete mitogenome of a non-albatross species of Procellariiformes. We used the MinION sequencing platform (Oxford Nanopore Technologies, UK, Jain et al. 2016) to obtain reads long enough to encompass the entire duplicated region, providing direct evidence of its existence.

Material and Methods

We focused on a single individual of *Puffinus lherminieri* from Martinique, caught and bled within the framework of a long-term demographic program on South Martinique, 14°25'02.8"N 60°49'53.7"W (see Precheur et al. 2016 for details on the study site and study species). Genomic DNA was extracted from the blood sample using the NucleoSpin® Tissue XS Kit (Macherey & Nagel, Düren, Germany). The sample was incubated overnight in 4 mg of Proteinase K. Purified genomic DNA was eluted twice in 50 µL of TE buffer pre-heated at 70°C. To ensure optimal PCR amplification, DNA quality and quantity were measured using 1% Agarose gel electrophoresis and Nanodrop 1000 spectrophotometry.

1. PCR amplification and Sanger sequencing of cox1

The mitochondrial genome was amplified using long-range PCR. First, we amplified 576 bp of c*ox1*. We chose *cox1* because it was expected to be located opposite the genes previously shown to be duplicated in other Procellariformes (namely *cob*, *nad6*, CR and several tRNAs), and we wanted the duplication to occur in the middle of the long-range PCR product, so as to be sure to sequence it. Bird blood is poorly concentrated in mitochondria and is likely to contain NUMTS

(Sorenson & Quinn 1998), i.e., nuclear copies of mitochondrial genes. Since the mitochondrial and nuclear genetic codes are not the same, these nuclear copies may be non-functional and may thus diverge from the mitogenome by genetic drift (Lopez et al. 1994). Co-sequencing of the nuclear and mitochondrial copies will thus lead to ambiguities in the sequences. To avoid such copies, we digested linear, nuclear genomic DNA prior to sequencing *cox1* (Sorenson & Quinn 1998), using ExonucleaseV (NEB-M0345S), according to the following protocol, modified from the manufacturer's instructions and from the protocol described in Jayaprakash et al. (2015): One nanogram (ng) of DNA sample was heated to 70°C to inactivate putative Proteinase K residual of the extraction protocol. Digestion was then carried out, adding the following to the sample in a 15 µL volume: 1X NEB4 Buffer, 1 mM ATP, 0.3 U of ExoV, 0.24 mg/mL of BSA. The mix was then heated to 37°C for 48h then to 70°C for 30 min to inactivate the exonuclease. This protocol allowed us to remove SNPs in numerous individuals of the same species in an upcoming population genetics paper (Torres et al. in prep).

Shearwater-specific primers for *cox1* were designed using Primer3 (Untergasser et al. 2012). PCRs were carried out in a total volume of 30 µL, using 1X Ex Taq Buffer (Mg2+ plus), 200 µM of dNTP, 0.8 µM of each primer, 0.025 U of TaKaRa Ex Taq® DNA Polymerase Hot-Start Version, and 60 ng of DNA extract. After an initial denaturation step of two min at 95°C, we ran 40 PCR cycles consisting of 1 min at 95°C, 1 min at 56°C and 1 min at 72°C. These cycles were followed by a 7-min final extension step at 72°C. PCR products were purified and sequenced on both DNA strands by Eurofins Genomics Munich, using the same PCR primers. Chromatograms were first visually inspected and edited using Sequencher v.5.4.1 (Gene Codes Corporation), and then sequences were assembled.

2. Long-Range amplification of the mitogenome

Based on *cox1* sequences from this individual, as well as others sampled from the same population (Torres et al. in prep), we designed three long-range primers conserved within this population, using Primer3 (Untergasser et al. 2012): Co1_test_L_221_1_LT (GCTCCTGCTTCTACTGTAGATGAGGCTAGTAGGAG) on the light strand, and Co1_test_H_344_1_EP (CGACCTAGCTATCTTCTCTCTTCACCTAGC) and Co1_test_H_296_1_EP (TTAGCCCATGCTGGAGCCTCAGTCGACCTAG) on the heavy strand.

Long-range PCR was carried out using TaKaRa LA Taq® DNA Polymerase Hot-Start Version, in a total volume of 25 µL, using 0.25 U of taq, 1x LA PCR Buffer II (Mg2+ plus), 0.21 mM of each dNTP, 0.125µM of each primer and 900 ng of purified (.e., free of nuclear DNA) mitochondrial DNA. After an initial denaturation step of one min at 94°C, we ran 35 PCR cycles consisting of 10s at 98°C, 30s at the primer-specific annealing temperature (with a temperature gradient from 60 to 72°C), and 17 to 17.5 min at 72°C. These cycles were followed by a 10-min final extension step at 72°C. PCR products were purified on agarose gel using the Monarch® DNA Gel Extraction Kit (New England Biolab) when multiple bands were visible. We followed the manufacturer's instruction, except that we added 1.5 times the volume of water and that we used 50 µL of elution buffer heated to 50°C.

3. Library preparation and sequencing

Given the small size differences expected between the two sets of amplified products (48 bp) all PCR products were pooled, and a final purification was performed using Beckman beads (Agencourt Bioscience). Quality control was performed using Qubit, Nanodrop and Fragment Analyzer. The Fragment Analyzer run revealed a peak of DNA concentration at around 12 kb,

although we were expecting 18kb amplicons. The smaller size of focus band of Fragment Analyzer could be explained by folding of the PCR products.

Samples were prepared for sequencing following the 1D DNA protocol selecting for long reads (SQK-LSK108, ONT). DNA repair, end repair, and A-tailing (M6630 and E7546, NEB) were performed on PCR products (3.8µg input) and each step was followed by a purification using Beckman beads. Adapters were ligated using the Blunt/TA ligase master mix (M0367, NEB). A 0.6X Beckman beads purification followed the ligation step, and 620 ng of library was loaded into the flowcell. DNA was not sheared so as to maximize sequencing read length. MinION sequencing was performed as per manufacturer's guidelines using R9.5 flowcells (FLO-MIN107, ONT), by MinKnow v1.7.10 (ONT), and runs extended for up to 48 h. The MUX scan reported 982 active pores.

4. Bioinformatics

Base calling was performed with the ONT Albacore command line tool (v1.2.4) and reads were outputted in the fastq format. We trimmed adapter sequences using Porechop (v0.2.1, Wick R. Porechop, available at: https://github.com/rrwick/Porechop). We used Nanofilt (v 1.1.3, De Coster et al. 2018, available at https://github.com/wdecoster/nanofilt) to filter reads for which the average quality score was less than 11.

Read assembly was performed using Canu (Koren et al. 2016) with the ng6 platform (Mariette et al. 2012). Because we were expecting a genome 16 kb long (18 kb long if the duplication was present) we choose to exclude from the analysis all the reads longer than 25 kb. By default, Canu also removes reads shorter than 1 kb. We expected that the duplicated region to be complex and thus difficult to assemble, so assembly was run with a target coverage setting of 100X. Similarly, as the obtained reads seemed to be of relatively low quality, we increased the

correction quality setting to "corMinCoverage=8." Resulting contigs and unitigs were annotated using MITOS (Bernt et al. 2013). We then compared them to the nearly-complete mitogenome of *Thalassarche melanophris* (Gibb et al. 2007) (AY158677.2) using Blast (MegaBLAST, nr database, E-value threshold: 10, identity threshold 85%) (Altschul et al. 1990).

Due to the presence of several stop-codons in the resulting contigs, we polished our Canu assembly using 100bp Illumina HiSeq 2500 reads previously obtained as a by-product of a separate, targeted enrichment project (Welch et al. in prep) on a separate individual from the same population (National Museum of Natural History, Smithsonian Institution, voucher: USNM 620721). The reads were aligned with the contigs using BWA (Li 2013), and then polished with Pilon (Walker et al. 2014). Illumina reads were mapped again on the resulting assembly sequence and drops in coverage were observed at some positions. These were resolved manually by correcting the assembly with the corresponding Illumina reads with the greatest depth. The resulting sequence was again annotated using MITOS.

Results & Discussion

We obtained 148 644 raw MinION reads (SRA accession SRS3196799). After length and quality filtering, we ran Canu with 12 764 reads. To optimize running time the assembly is constructed so that 100x of the 18 expected kb is covered. This is why only 88 reads were retained for the final assembly; these had a total cumulative length of 1 748 405 bp. The median length of the retained reads was 21 203 bp. Of these 88 reads, 74 (84%) were used to assemble a contig 21 344 bp long (hereafter named Ct1; average coverage 57X, min 9X, max 74X) and 12 reads (14%) were used to assemble a contig of 18 884 bp (hereafter named Ct2; average coverage 10X, min 5X, max 12X). The two remaining orphan reads were removed from further analyses,

as they corresponded to short mitochondrial sequences, slightly divergent from Ct1 and Ct2. We mapped the quality-filtered 12 764 MinION reads onto these two contigs: 88% mapped to Ct1 and 92% to Ct2. The average coverage was high, with more than 3 000 X for the two contigs (Tab. 2). We observed, however, that coverage of the first 8.4 kb of the two genomes was almost ten times lower than that of the rest of the genome (Fig. 1a, Fig. 2a, Tab. 2). This difference in coverage was not visible when mapping only the 88 reads used for the assembly, or when mapping the Illumina reads (see below). These two regions presented a similar GC content on the two assemblies (Tab. 2), suggesting that the first half was not more complex than the second one. We therefore suggest that this difference of coverage was due to a difficulty in sequencing encountered by the MinION device (e.g., due to secondary structures formed by mitochondrial DNA).

We mapped 277 693 Illumina reads on the two contigs (SRA acc. SRP141134), obtaining an average coverage of more than 1 300 X for both (Fig. 1b, Fig. 2b, Tab. 3). Using these data, we corrected 150 (on Ct1) and 177 (on Ct2) local SNPs on the assemblies. The two resulting genomes (GenBank acc. MH206162 and MH206163) were 21 144 bp long and 19 004 bp long, respectively, and consisted of 14 and 13 protein coding genes, respectively, 25 tRNA genes and 2 rRNA genes (Fig. 1c, Fig. 2c, Supplementary Tab. 1, Supplementary Tab. 2). The nucleotide composition of the complete genome was 32.0% A, 29.9% C, 11.8% G and 26.3% T for Ct1 and 31.1% A, 29.8% C, 12.8% G and 26.3% T for Ct2, as expected in AT-rich mitochondrial genomes (see Saccone et al. 1999). Most of the genes are encoded on the light strand, with only *nad6* and eight tRNA genes (Gln, Ala, Asn, Cys, Try, Glu, Pro and Ser) encoded on the heavy strand.

All protein-coding genes started with typical ATN codons, except for *cox2* and *nad5*, which began with the GTG codon. All protein coding genes finished with the TAA stop-codon,

except for *nad4*, *nad5* and *nad1*, which ended with AGA, *nad2* and *nad6*, which ended with TAG, and *cox1*, which ended with AGG (Supplementary Tab. 1 and Supplementary Tab. 2). We observed a supplementary base in the gene *nad3* that implies a translational frameshift (Mindell et al. 1998). This had already been observed in mitogenomes of several bird species, including Procellariiformes (e.g. Watanabe et al. 2006; Gibb et al. 2013).

Ct1 and Ct2 were different in length and in composition. Although we cannot exclude that the DNA sample was contaminated by another one, we have strong evidence to contradict this hypothesis (see Supplementary Material 1). We observed that a duplicated region was present on Ct1, whereas Ct2 did not include a duplication. We used the contigs obtained with the MinION reads to study local divergences between the two mitogenomes. We corrected the contigs for indels but not for substitutions, with Illumina reads, using Pilon (Walker et al. 2014). We found 52 SNPs between the two contigs, only five of which were transversions. The distribution of the SNPs was not homogenous along the genome because *nad-4*, *nad-6*, the control region and *rnL* contained 35 (67%) of them. Of the 52 SNPs, 37 were located in coding regions, and 43% were on the third position of the codon and were synonymous mutations. We inferred the consensus sequence of the two contigs, using ambiguity codes where the sequences diverged. We mapped the 12 764 MinION reads on this consensus sequence and observed that 85% of the reads were consistent with Ct1, whereas 15% of the reads were consistent with Ct2. The distribution of the SNPs was not homogeneous along the genome and the divergence among the contigs, consistent with all the reads. These characteristics made it unlikely for sequencing error to be the sole explanation for the divergence between the two contigs. Based on these results, we suggest that the *Puffinus lherminieri* individual used for MinION sequencing showed mitochondrial heteroplasmy. Heteroplasmy has already been observed in birds (e.g. Mundy et al. 1996; Moum & Bakke 2001; Gandolfi et al. 2017). Because of likely PCR artefacts, such as preferential

amplification of one mitogenome over the other, the proportion of MinION reads did not necessarily reflect the proportion of the two mitogenomes in the individual. When we mapped all the Illumina reads on these two contigs, we did not see any divergence among the reads, suggesting that this heteroplasmy was not present in the individual sequenced by Illumina.

The duplication in Ct1 consisted of two identical copies of *nad6*, two identical copies of the tRNAs Phe, Trp and Cys, and two dissimilar copies of the CR (which we call CR1 and CR2). Contrary to previous results for albatrosses (Abbott et al. 2005, Gibb et al. 2007, Eda et al. 2010, Lounsberry et al. 2015), *cob* was not duplicated CR1 and CR2 included a 1 270 bp region in common; 24 mutations separated CR1 from CR2. This overlap region is followed in CR1 by 229 supplementary bases and in CR2 by 2 033 supplementary bases. These two supplementary sequences did not align to each other. The single CR found in Ct2, which we will call CR3, was more similar to CR1 than to CR2 in the overlap region but included a 2 250 bp supplementary region that aligned with the supplementary region of CR2.

CR2 (on Ct1) and CR3 (on Ct2) were followed by a 2 kb-long stretch of DNA (hereafter called RR, for repeat region) (Fig. 1, Fig. 2) composed of 90 bp modules repeated 19 times. During polishing, Illumina reads mapped onto the beginning of the RR, but no read overlapped CR3/CR2 and the RR, nor did they overlap the RR and 12S; hence we have no proof that the RR was effectively present in the mitogenome of *P. lherminieri* individual USNM 620721 (Illumina). Moreover, the RR was poorly covered by the Illumina reads (Tab. 3), and in Ct1 this region had a lower GC content (32%) than the rest of the genome (average of 42%). This could explain why this region was more difficult to sequence with the Illumina technology. MinION reads were discordant in this particular region. When the RR was manually removed from the genome, no read linked CR3/CR2 and 12S. This means that no MinION read covered the entire mitochondrial genome if the RR was not present. Moreover, no significant BLAST match was

found for this region (MegaBLAST & BLASTn, nr database, E-value threshold: 10, identity threshold 85%). Therefore, two hypotheses may explain the presence of this repeated sequence: (i) the RR is biologically present in the mitochondrial genome of this individual of *Puffinus lherminieri*, and potentially in other species of Procellariiformes, but was never sequenced until now. The large number of repetitions implied that MinION sequencing was difficult and led to differences among reads. Illumina sequencing was difficult in this region due to the high AT content and so, no Illumina read linked the RR to the rest of the mitogenome. Alternatively, (ii) the RR may be an artefact, due for example to chimeric amplification in the early stages of the PCR, containing parts of the mitochondrial genome and a nuclear sequence not catalogued on Genbank.

The duplicated region of *Puffinus lherminieri* was similar to all the duplicated regions observed in albatross species. The two Ct1 copies of *nad6* and the tRNA were identical, and the two control regions differed by several bases in the first 100 bp. The only major difference between shearwater and albatross mitogenome structures was that no duplication of *cob* was observed in the genome of *Puffinus lherminieri*. We also found evidence for two different mitogenomes co-existing in the same individual. These two mitogenomes differ by point mutations and gene duplication. Because albatross mitogenomes were sequenced using short fragments, heteroplasmy may have been harder to detect.

To conclude, we have provided direct evidence for a duplicated region in the mitogenome of *Puffinus lherminieri*. A similar duplication had already been observed in several albatross species (Abbott et al. 2005; Gibb et al. 2007; Eda et al. 2010; Lounsberry et al. 2015), quite phylogenetically distant from the shearwaters (Welch et al. in prep). This suggests that the duplication may be widespread within the Procellariiformes. The fact that the composition of the duplication is different between albatrosses and shearwaters suggests that this region has evolved

during history of the Procellariiformes. At least one event of deletion or addition of the copy of *cob* has occurred between diversification of albatrosses and shearwaters, and we know that other species show different versions of this duplication (Gibb et al. 2013). Investigating mitogenomes from more species is needed to better understand the evolutionary history of the mitochondrial genome of the Procellariiformes.

Acknowledgements

We thank the University of La Rochelle (ULR) for financial support, and the ULR molecular core facility for lab support. We thank Helen James and the National Museum of Natural History, Smithsonian Institution, for access to tissue sample USNM 620721; UCE data for this individual was originally gathered for an unpublished phylogenetics study of Procellariiformes conducted by AJW, RTC, Helen James, and VB. We are grateful to the bioinformatics platform Toulouse Midi-Pyrenees (Bioinfo GenoToul) for providing help and computing. We also thank C. Precheur, the Parc Naturel Regional de la Martinique and the DEAL Martinique for help and funding of the research program on the shearwater. Salary of LT is covered by a grant from the University of La Rochelle. Any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

Declaration of Interest statement

The authors report no conflicts of interest

Funding details

This work is covered by a grant from the University of La Rochelle and funding from the National Museum of Natural History, Smithsonian Institution.

References

Abbott CL, Double MC, Trueman JWH, Robinson A, Cockburn A. 2005. An unusual source of apparent mitochondrial heteroplasmy: Duplicate mitochondrial control regions in *Thalassarche* albatrosses. *Mol. Ecol.* 14(11):3605–13

Alderman R, Double MC, Valencia J, Gales RP. 2005. Genetic affinities of newly sampled populations of Wandering and Black-browed Albatross. *Emu*. 105:169–79

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *J. Mol. Biol.* 215(3):403–10

Austin JJ, Bretagnolle V, Pasquet E. 2004. A global molecular phylogeny of the small *Puffinus* shearwaters and implications for systematics of the Little-Audubon's Shearwater complex. *Auk*. 121(3):647–864

Avise JC, Arnold J, Ball RM, Bermingham E, Lamb T, et al. 1987. Intraspecific Phylogeography; the Mitochondrial DNA Bridge Between Population Genetics and Systematics. *Annu. Rev. Ecol. Syst*. 18:489–522

Berg T, Moum T, Johansen S. 1995. Variable numbers of simple tandem repeats make birds of the order Ciconiiformes heteroplasmic in their mitochondrial genomes. *Curr. Genet*. 27(3):257–62

Bernt M, Donath A, Jühling F, Externbrink F, Florentz C, et al. 2013. MITOS: Improved de novo metazoan mitochondrial genome annotation. *Mol. Phylogenet. Evol.* 69(2):313–19

Boore JL. 1999. Animal mitochondrial genomes. *Nucleic Acids Res.* 27(8):1767–80

Brown WM. 1985. The mitochondrial genome of animals. *Mol. Evol. Genet*. 95–130

Burg TM, Bird H, Lait L, de M. 2014. Colonization pathways of the northeast Atlantic by northern fulmars: A test of James Fisher's "out of Iceland" hypothesis using museum collections. *J. Avian Biol.* 45(3):209–18

Burg TM, Croxall JP. 2001. Global relationships amongst black-browed and grey-headed albatrosses: analysis of population structure using mitochondrial DNA and microsatellites. *Mol. Ecol.* 10:2647–60

Cagnon C, Lauga B, Hémery G, Mouchès C. 2004. Phylogeographic differentiation of storm petrels (*Hydrobates pelagicus*) based on cytochrome b mitochondrial DNA variation. *Mar.*

Biol. 145(6):1257–64

De Coster, W., D'Hert, S., Schultz, D. T., Cruts, M., & Van Broeckhoven, C. (2018). NanoPack: visualizing and processing long read sequencing data. *bioRxiv*, 237180

Eda M, Kuro-o M, Higuchi H, Hasegawa H, Koike H. 2010. Mosaic gene conversion after a tandem duplication of mtDNA sequence in Diomedeidae (albatrosses). *Genes Genet. Syst.* 85(2):129–39

Gandolfi A, Crestanello B, Fagotti A, Simoncelli F, Chiesa S, et al. 2017. New Evidences of Mitochondrial DNA Heteroplasmy by Putative Paternal Leakage between the Rock Partridge (*Alectoris graeca*) and the Chukar Partridge (*Alectoris chukar*). *PLoS One*. 12(1):4–11

Gangloff B, Zino F, Shirihai H, González-Solís J, Couloux A, et al. 2013. The evolution of north-east Atlantic gadfly petrels using statistical phylogeography. *Mol. Ecol.* 22(JANUARY):495–507

Gibb GC, Kardailsky O, Kimball RT, Braun EL, Penny D. 2007. Mitochondrial genomes and avian phylogeny: Complex characters and resolvability without explosive radiations. *Mol. Biol. Evol.* 24(1):269–80

Gibb GC, Kennedy M, Penny D. 2013. Beyond phylogeny: Pelecaniform and ciconiiform birds, and long-term niche stability. *Mol. Phylogenet. Evol.* 68(2):229–38

Harrison P. 2000. *Seabirds: An Identification Guide (No. QL 673. H38 2000)*. Houghton Mifflin CO., New York

Jain M, Olsen HE, Paten B, Akeson M. 2016. The Oxford Nanopore MinION: delivery of nanopore sequencing to the genomics community. *Genome Biol.* 17(1):256

Jayaprakash AD, Benson EK, Gone S, Liang R, Shim J, et al. 2015. Stable heteroplasmy at the single-cell level is facilitated by intercellular exchange of mtDNA. *Nucleic Acids Res.* 43(4):2177–87

Jesus J, Menezes D, Gomes S, Oliveira P, Nogales M, Brehm A. 2009. Phylogenetic relationships of gadfly petrels *Pterodroma spp*. from the Northeastern Atlantic Ocean: molecular evidence for specific status of Bugio and Cape Verde petrels and implications for

conservation. *Bird Conserv. Int.* 19(3):199

Katoh K, Misawa K, Kuma K, Miyata T. 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* 30(14):3059–66

Kerr KCR, Dove CJ. 2013. Delimiting shades of gray: phylogeography of the Northern Fulmar , *Fulmarus glacialis*. . 1915–30

Koren S, Walenz BP, Berlin K, Miller JR, Bergman NH, Phillippy AM. 2016. Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome Res.* 27(5):722–36

Kuro-o M, Yonekawa H, Saito S, Eda M, Higuchi H, et al. 2010. Unexpectedly high genetic diversity of mtDNA control region through severe bottleneck in vulnerable albatross Phoebastria albatrus. *Conserv. Genet.* 11(1):127–37

Lawrence HA, Lyver POB, Gleeson DM. 2014. Genetic panmixia in New Zealand's Grey-faced Petrel: Implications for conservation and restoration. *Emu.* 114(3):249–58

Lawrence HA, Taylor GA, Millar CD, Lambert DM. 2008. High mitochondrial and nuclear genetic diversity in one of the world's most endangered seabirds, the Chatham Island Taiko (*Pterodroma magentae*). *Conserv. Genet.* 9(5):1293–1301

Li H. 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv Prepr. arXiv.* 0(0):1–3

Lopez J V., Yuhki N, Masuda R, Modi W, O'Brien SJ. 1994. Numt, a recent transfer and tandem amplification of mitochondrial DNA to the nuclear genome of the domestic cat. *J. Mol. Evol.* 39(2):174–90

Lounsberry ZT, Brown SK, Collins PW, Henry RW, Newsome SD, Sacks BN. 2015. Next-generation sequencing workflow for assembly of nonmodel mitogenomes exemplified with North Pacific albatrosses (*Phoebastria spp.*). *Mol. Ecol. Resour.* 15(4):893–902

Mariette J, Escudié F, Allias N, Salin G, Noirot C, et al. 2012. NG6: Integrated next generation sequencing storage and processing environment. *BMC Genomics.* 13(1):462

Mindell DP, Sorenson MD, Dimcheff DE. 1998. An Extra Nucleotide Is Not Translated in Mitochondrial ND3 of Some Birds and Turtles. *Mol. Biol. Evol.* 15(11):1568–71

Moritz C, Dowling TE, Brown WM. 1987. Evolution of Animal Mitochondrial DNA: Relevance for Population Biology and Systematics. *Annu. Rev. Ecol. Syst.* 18:269–92

Moum T, Bakke I. 2001. Mitochondrial control region structure and single site heteroplasmy in the razorbill (*Alca torda* ; Aves). *Curr. Genet.* 39:198–203

Mundy NI, Winchell CS, Woodruff DS. 1996. Tandem Repeats and Heteroplasmy in the Mitochondrial DNA Control Region of the Loggerhead Shrike (*Lanius ludovicianus*). *J. Hered.* 87(1):1–6

Peck DR. 2006. *Local adaptation in the wedge-tailed shearwater (*Puffinus pacificus*). PhD thesis, James Cook University.*

Precheur C, Barbraud C, Martail F, Mian M, Nicolas J, et al. 2016. Some like it hot : effect of environment on population dynamics of a small tropical seabird in the Caribbean region. *Ecosphere*. 7(10):1–18

Prum RO, Berv JS, Dornburg A, Field DJ, Townsend JP, et al. 2015. A comprehensive phylogeny of birds (Aves) using targeted next-generation DNA sequencing. *Nature*. 526(7574):569–73

Rains D, Weimerskirch H, Burg TM. 2011. Piecing together the global population puzzle of wandering albatrosses: Genetic analysis of the Amsterdam albatross Diomedea amsterdamensis. *J. Avian Biol.* 42(1):69–79

Ramírez O, Gómez-Díaz E, Olalde I, Illera JC, Rando JC, et al. 2013. Population connectivity buffers genetic diversity loss in a seabird. *Front. Zool.* 10(1):28

Saccone C, De Giorgi C, Gissi C, Pesole G, Reyes A. 1999. Evolutionary genomics in Metazoa: The mitochondrial DNA as a model system. *Gene*. 238(1):195–209

Sammler S, Bleidorn C, Tiedemann R. 2011. Full mitochondrial genome sequences of two endemic Philippine hornbill species(Aves: Bucerotidae) provide evidence for pervasive mitochondrial DNA recombination. *BMC Genomics*. 12(1):35

Slack KE, Jones CM, Ando T, Harrison GL, Fordyce RE, et al. 2006. Early penguin fossils, plus mitochondrial genomes, calibrate avian evolution. *Mol. Biol. Evol.* 23(6):1144–55

Smith AL, Monteiro L, Hasegawa O, Friesen VL. 2007a. Global phylogeography of the band-

rumped storm-petrel (*Oceanodroma castro*; Procellariiformes: Hydrobatidae). *Mol. Phylogenet. Evol.* 43(3):755–73

Smith AL, Monteiro L, Hasegawa O, Friesen VL. 2007b. Global phylogeography of the band-rumped storm-petrel (*Oceanodroma castro*; Procellariiformes: Hydrobatidae). *Mol. Phylogenet. Evol.* 43(3):755–73

Sorenson MD, Quinn TW. 1998. NUMTS: a challenge for avian systematics and population biology. *Auk*. 115:214–21

Stamatakis A. 2014. RAxML Version 8 : A tool for Phylogenetic Analysis and Post-Analysis of Large Phylogenies. . 2010–11

Tsaousis AD, Martin DP, Ladoukakis ED, Posada D, Zouros E. 2005. Widespread recombination in published animal mtDNA sequences. *Mol. Biol. Evol.* 22(4):925–33

Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, et al. 2012. Primer3-new capabilities and interfaces. *Nucleic Acids Res.* 40(15):1–12

Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, et al. 2014. Pilon: An integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One*. 9(11):

Watanabe M, Nikaido M, Tsuda TT, Kobayashi T, Mindell D, et al. 2006. New candidate species most closely related to penguins. *Gene*. 378(1–2):65–73

Welch AJ, Olson SL, Fleischer RC. 2014. Phylogenetic relationships of the extinct St Helena petrel, *Pterodroma rupinarum* Olson, 1975 (Procellariiformes: Procellariidae), based on ancient DNA. *Zool. J. Linn. Soc.* 170(3):494–505

Tables

Tab. 1 Previously known mitochondrial duplications in Procellariiformes.

NA: information not available, CR: Control Region

Family	Species	*cob* duplicated	*nad6* duplicated	CR duplicated	Object of the study	Genbank accession number	Sequencing method	Study
Diomedeidae	*Diomedea amsterdamensis, D. exulans*	NA	NA	yes	CR		PCR and Sanger-sequencing	Rains et al. 2011
Diomedeidae	*Diomedea chrysostoma*	NA	NA	NA	nearly complete mitogenome	AP009193.1	PCR, primer-walking and shotgun sequencing	Watanabe et al. 2006
Diomedeidae	*Phoebastria albatrus*	NA	NA	yes	CR		PCR and Sanger sequencing	Kuro-o et al. 2010
Diomedeidae	*Phoebastria albatrus, Ph. immutabilis, Ph. nigripes*	in part	yes	yes	whole duplicated region	AB276044: AB276051	PCR, primer-walking and Sanger sequencing	Eda et al. 2010
Diomedeidae	*Phoebastria albatrus, Ph. immutabilis, Ph. nigripes*	in part	yes	yes	complete mitogenome	KJ735512.1: KJ735514.1	Illumina sequencing, PCR and Sanger sequencing	Lounsberry et al. 2015
Diomedeidae	*Thalassarche cauta*	in part	yes	yes	whole duplicated region		Restriction digest map, PCR, primer-walking, Sanger sequencing	Abbott et al. 2005
Diomedeidae	*Thalassarche melanophris*	in part	yes	yes	complete mitogenome	AY158677.2	Re-check from Slack et al. 2006	Gibb et al. 2007
Hydrobatidae	*Oceanodroma castro*	NA	NA	yes	CR			Smith et al. 2007
Procellariidae	*Fulmarus glacialis*	NA	NA	yes	CR		PCR and Sanger sequencing	Burg et al. 2014
Procellaridae	*Procellaria cinerea*	NA	NA	NA	nearly complete mitogenome	AP009191.1	PCR, primer-walking and shotgun sequencing	Watanabe et al. 2006
Procellariidae	*Pterodroma brevirostris*	NA	NA	NA	complete mitogenome	AY158678.1	PCR and Primer-walking, Sanger sequencing	Slack et al. 2006
Procellariidae	*Pterodroma macroptera gouldi*	NA	NA	yes	CR		PCR and Illumina sequencing	Lawrence et al. 2014
Procellariidae	*Pterodroma magentae*	NA	NA	yes	CR		PCR and Sanger sequencing	Lawrence et al. 2008

Tab. 2 Coverage (X) of mapping of the 12 764 MinION reads used by Canu for the assembly, and percentage of GC, for the two contigs. The "first part" consists of the first 8 428 bp of Ct1 and the first 8 447 bp of Ct2, and the "second part" consists of the last 12 716 bp of Ct1 and the last 10 557 bp of Ct2.

Sequence	Average coverage	Minimum coverage	Maximum coverage	% of covered bases	% of GC
Ct1	3 189	433	5 927	100	41,76
Ct2	3 523	433	9 067	100	42,66
Ct1 first part	560	433	683	100	43,36
Ct2 first part	560	433	683	100	43,36
Ct1 second part	4 945	3 812	5 927	100	40,70
Ct2 second part	5 919	4 229	9 067	100	42,11

Tab. 3 Coverage (X) of mapping of the 277 693 Illumina reads and percentage of GC of the two genomes. "Ct without RR" is the genome sequence in which the RR sequence was manually deleted.

Sequence	Average coverage	Minimum coverage	Maximum coverage	% of covered bases	% of GC
Ct1	1 327	0	6 971	95	41.8
Ct2	1 386	0	4 554	93	42.7
RR in Ct1	575	0	6 971	41	32.3
RR in Ct2	40	0	745	12	42.7
Ct1 without RR	1 395	465	2 449	100	42.7
Ct2 without RR	1 547	574	4 554	100	42.8

Figures

Figure 1.

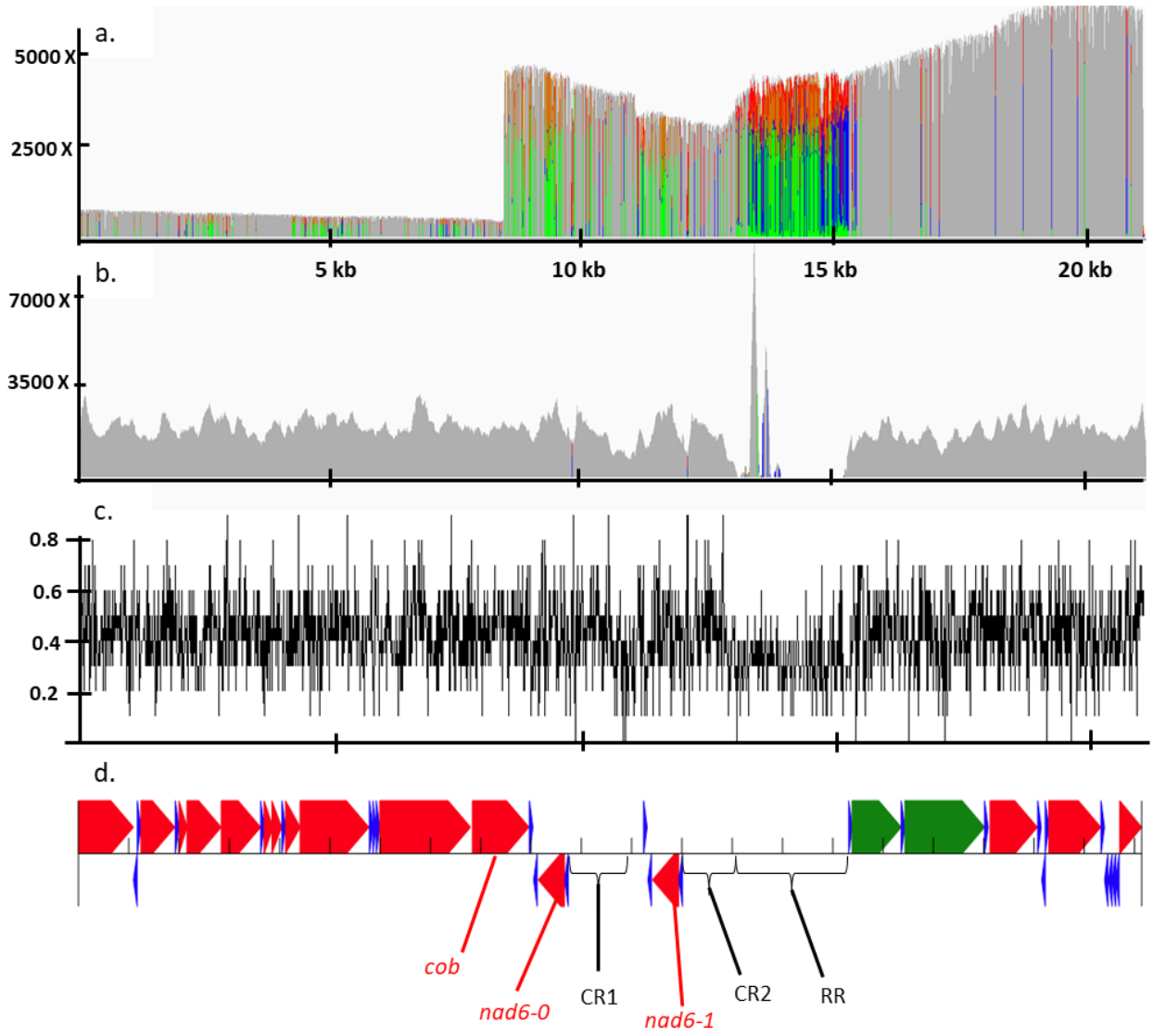


Figure 2.

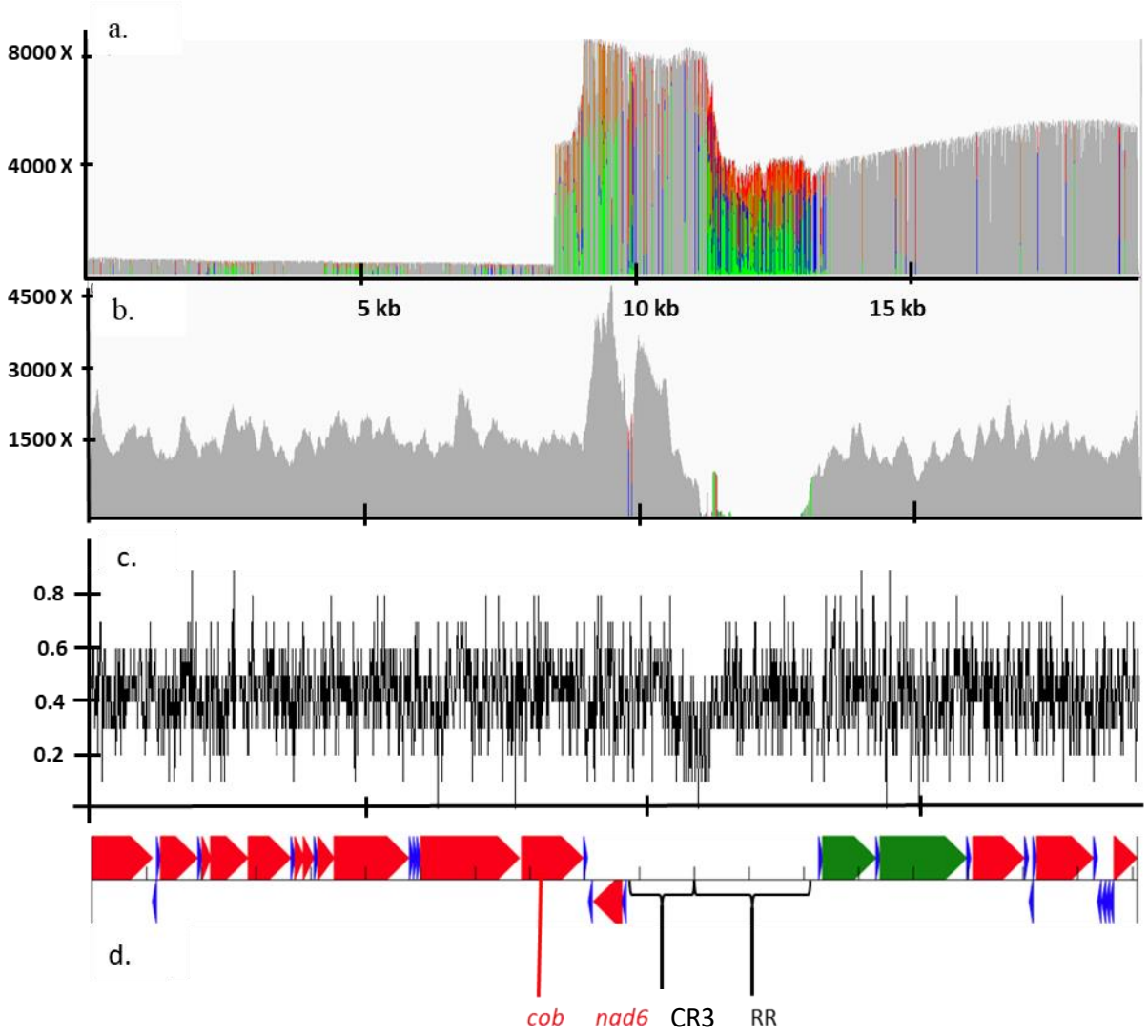

a.
8000 X
4000 X
4500 X
b.
5 kb
10 kb
15 kb
3000 X
1500 X
c.
0.8
0.6
0.4
0.2
d.
cob
nad6
CR3
RR

Figure Captions

Figure 1. Coverage of mapped MinION and Illumina reads and MITOS annotation for the longer mitogenome (Ct1). a. Coverage of mapping of the 12 764 MinION reads used by Canu for the assembly. Colored bars represent SNPs among reads, gray bars represent consensus bases among all the reads. b. Coverage of mapping of the 277 693 Illumina reads. c. GC content. d. MITOS annotation. Red, green and blue arrows represent protein-coding genes, ribosomal genes, and tRNAs, respectively.

Figure 2. Coverage of mapped MinION and Illumina reads and MITOS annotation for the shorter mitogenome (Ct2). a. Coverage of mapping of the 12 764 MinION reads used by Canu for the assembly. Colored bars represent SNPs among reads, gray bars represent consensus bases among all the reads. b. Coverage of mapping of the 277 693 Illumina reads. c. GC content. d. MITOS annotation. Red, green and blue arrows represent protein-coding genes, ribosomal genes, and tRNAs, respectively.

Supplementary Tab. 1 Composition of Ct1

Name	Start	Stop	Strand	Length	Start-codon	Stop-codon
cox1_b	1	1092	+	1091		AGG
trnS2(tca)	1096	1169	-	73		
trnD(gac)	1172	1240	+	68		
cox2	1242	1916	+	674	GTG	TAA
trnK(aaa)	1927	1998	+	71		
atp8	2000	2161	+	161	ATG	TAA
atp6	2158	2838	+	680	ATG	TAA
cox3	2841	3623	+	782	ATG	TAA
trnG(gga)	3625	3693	+	68		
nad3_a	3694	3867	+	173	ATT	
nad3_b	3848	4042	+	194		TAA
trnR(cga)	4048	4116	+	68		
nad4l	4118	4411	+	293	ATG	TAA
nad4	4408	5775	+	1367	ATG	AGA
trnH(cac)	5786	5855	+	69		
trnS1(agc)	5856	5922	+	66		
trnL1(cta)	5922	5992	+	70		
nad5	5993	7795	+	1802	GTG	AGA
cob	7834	8958	+	1124	ATC	TAA
trnT(aca)	8968	9037	+	69		
trnP(cca)	9054	9123	-	69		
nad6-0	9147	9665	-	518	ATG	TAG
trnE(gaa)	9669	9741	-	72		
trnT(aca)	11240	11309	+	69		
trnP(cca)	11326	11395	-	69		
nad6-1	11419	11937	-	518	ATG	TAG
trnE(gaa)	11941	12013	-	72		
trnF(ttc)	15317	15386	+	69		
rrnS	15386	16361	+	975		
trnV(gta)	16361	16433	+	72		
rrnL	16434	18018	+	1584		
trnL2(tta)	18018	18091	+	73		
nad1	18129	19067	+	938	ATC	AGA
trnI(atc)	19075	19146	+	71		
trnQ(caa)	19156	19226	-	70		
trnM(atg)	19226	19294	+	68		
nad2	19295	20329	+	1034	ATG	TAG
trnW(tga)	20334	20403	+	69		
trnA(gca)	20405	20473	-	68		

trnN(aac)	20484	20556	-	72		
trnC(tgc)	20559	20625	-	66		
trnY(tac)	20626	20695	-	69		
cox1_a	20706	21143	+	437	ATC	

Supplementary Tab. 2 Composition of Ct2

Name	Start	Stop	Strand	Length	Start-codon	Stop-codon
cox1_b	2	1105	+	1103		AGG
trnS2(tca)	1109	1182	-	73		
trnD(gac)	1185	1253	+	68		
cox2	1255	1929	+	674	GTG	TAA
trnK(aaa)	1940	2011	+	71		
atp8	2013	2174	+	161	ATG	TAA
atp6	2171	2851	+	680	ATG	TAA
cox3	2854	3636	+	782	ATG	TAA
trnG(gga)	3638	3706	+	68		
nad3_a	3707	3880	+	173	ATT	
nad3_b	3861	4055	+	194		TAA
trnR(cga)	4061	4129	+	68		
nad4l	4131	4424	+	293	ATG	TAA
nad4	4421	5788	+	1367	ATG	AGA
trnH(cac)	5799	5868	+	69		
trnS1(agc)	5869	5935	+	66		
trnL1(cta)	5935	6005	+	70		
nad5	6006	7808	+	1802	GTG	AGA
cob	7847	8971	+	1124	ATC	TAA
trnT(aca)	8981	9050	+	69		
trnP(cca)	9067	9136	-	69		
nad6	9160	9678	-	518	ATG	TAG
trnE(gaa)	9682	9754	-	72		
trnF(ttc)	13272	13341	+	69		
rrnS	13341	14316	+	975		
trnV(gta)	14316	14388	+	72		
rrnL	14389	15973	+	1584		
trnL2(tta)	15973	16046	+	73		
nad1	16084	17022	+	938	ATC	AGA
trnI(atc)	17030	17101	+	71		
trnQ(caa)	17111	17181	-	70		
trnM(atg)	17181	17249	+	68		
nad2	17250	18284	+	1034	ATG	TAG
trnW(tga)	18289	18358	+	69		
trnA(gca)	18360	18428	-	68		
trnN(aac)	18439	18511	-	72		
trnC(tgc)	18514	18580	-	66		
trnY(tac)	18581	18650	-	69		
cox1_a	18661	19083	+	422	ATC	

Supplementary Material 1:

There are two explanations for the fact that the assembly of the MinION reads resulted in two contigs: either the DNA library was contaminated by another individual, so that the two contigs correspond to the mitochondrial genomes of two individuals, or two different mitogenomes were present in the one individual sequenced, i.e., heteroplasmy. We cannot rule out contamination, but we have strong evidence to contradict this hypothesis.

First, a Sanger sequence of *cox1* of the same individual groups with the MinION sequences. This sequence was obtained as part of a population genetics study of 175 individuals (Torres et al. in prep), representing all populations of *Puffinus lherminieri* and three populations of the closely related species *P. bailloni* (Genbank acc. MH383332-MH383506). The *cox1* sequences of the two MinION contigs were aligned with these 175 sequences using MAFFT (Katoh et al. 2002) and truncated to the 557 bp aligned for the other sequences. A phylogenetic tree was inferred using MrBayes and 500 million generations, with *P. yelkouan* (Genbank AY567884.1) as the outgroup. After deletion of 10% burn-in, convergence of the run was assessed using Tracer. The *cox1* sequences of the two MinION contigs were part of the same clade as the sequences of the same lineage obtained by Sanger sequencing (Fig. S1), with a posterior probability of 0.80. The sequence from Ct1 was identical to the sequence from the individual of origin. The sequence of Ct2 was different to the original individual from 1 SNP.

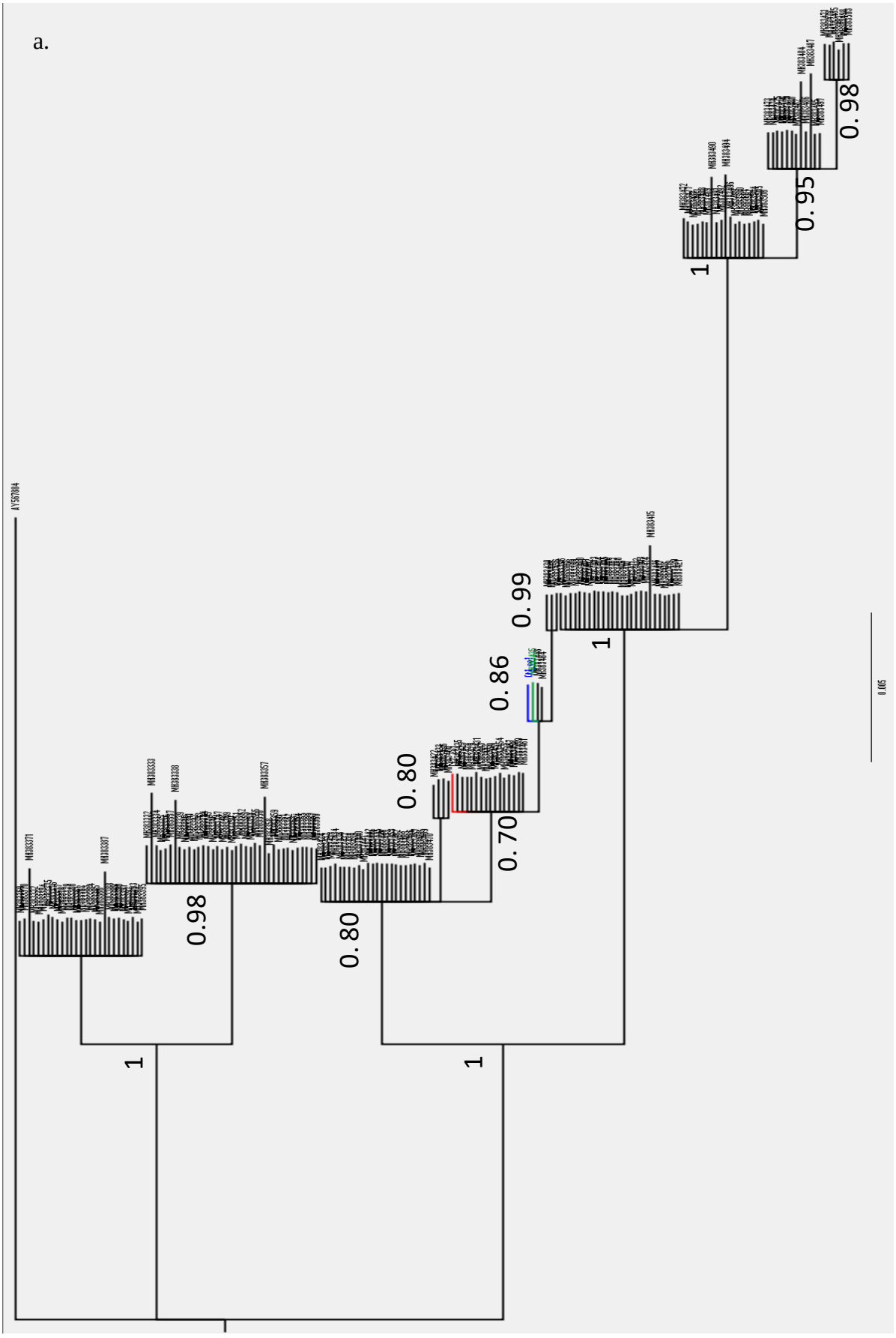
Second, control region sequences from Ct1 and Ct2 are identical to those obtained using Sanger sequencing. Partial control region sequences (313 bp) of 224 sequences of the control region from the same sampling locations are extremely polymorphic, consisting of different haplotypes for most individuals. The two control regions of Ct1 and the one of Ct2 were aligned and truncated to correspond to these Sanger sequences. These three sequences are both identical among themselves and to the sequence of the individual obtained by Sanger.

Thus, we compared two parts of the two mitochondrial genomes with sequences of other individuals, presenting numerous SNPs among individuals. These analyses show that both contigs are phylogenetically very close to the target individual. If the sample was contaminated it would be by a nearly identical sample phylogenetically, which is very unlikely.

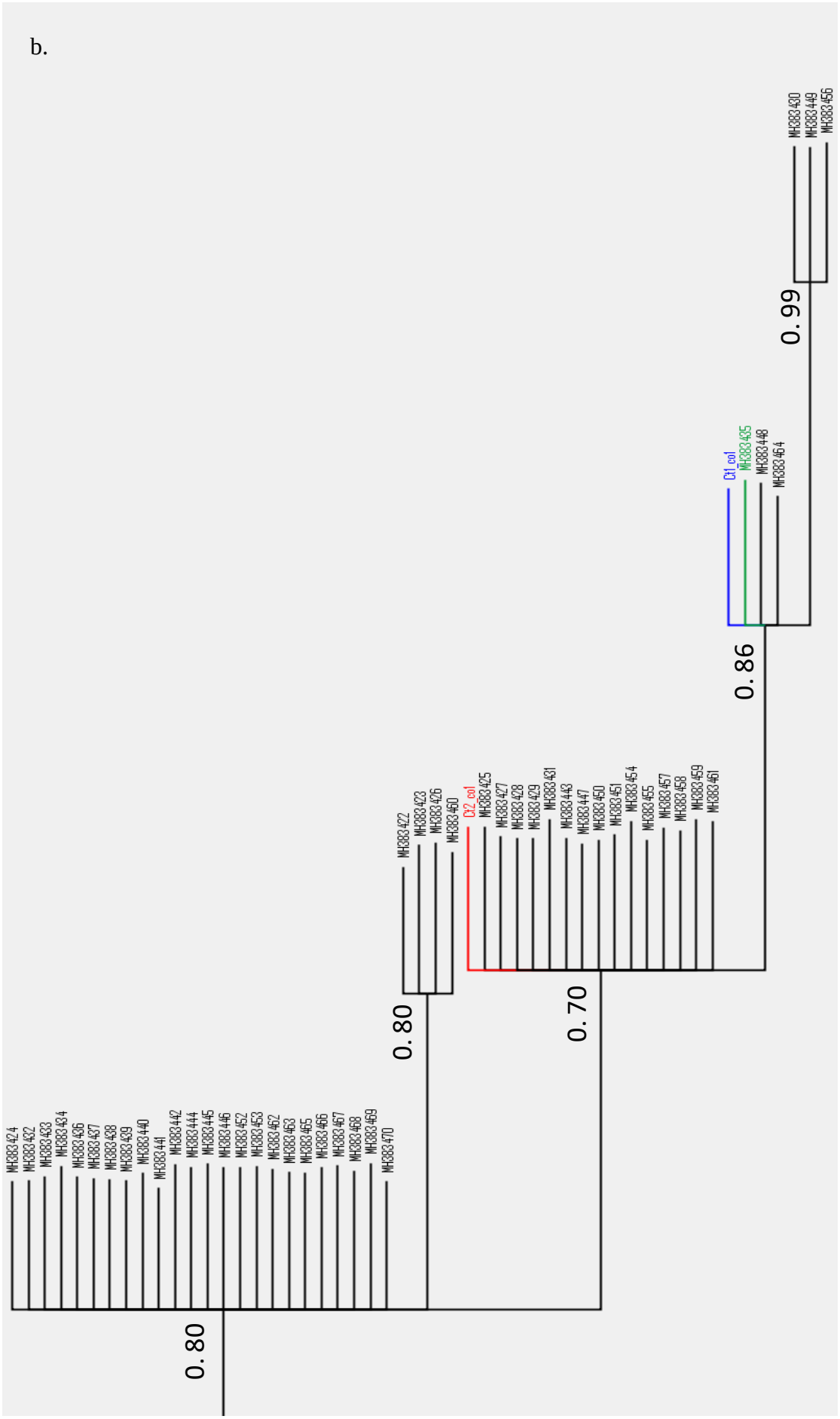
Moreover, we know that heteroplasmy occurs in birds (Mundy et al. 1996, Moum & Bakke 2001, Gandolfi et al. 2017), and we have no reason to think that it is not present in Procellariiformes. Hence several different mitogenomes can be present in the same individual. A cross-contaminated sample would bear more likely four different mitogenomes instead of two. So the assembly of two different mitogenomes is not an evidence of cross-contamination.

Therefore, we believe that the presence of two different contigs in our study is due to heteroplasmy. However, if the contamination is real, it would mean that one of the two individuals did not have the mitochondrial duplication (see Results & Discussion), and would suggest that the mitochondrial duplication is not present in every individual of this species. Therefore the evolutionary scenario of the duplicated region would be even more complicated than that suggested by the heteroplasmy hypothesis.

Supplementary Figure 1. a. Gene tree obtained from Sanger *cox1* sequences and the two *cox1* from the two MinION contigs. b. Zoom on the clade containing the two contig sequences. The sequence from Ct1 is highlighted in blue, the one from Ct2 in red, and the one from the same individual sequenced by Sanger in green. Posterior probabilities superior to 0.70 are shown.


a.
AY507084
MH383371
MH383387
MH383333
MH383330
MH383357
MH383345
MH383400
MH383494
MH383464
MH383407
0.98
0.80
0.80
0.70
0.86
0.99
1
1
1
0.95
0.98
0.005

b.


MH383424
MH383432
MH383433
MH383434
MH383436
MH383437
MH383438
MH383439
MH383440
MH383441
MH383442
MH383444
MH383445
MH383446
MH383452
MH383453
MH383462
MH383463
MH383465
MH383466
MH383467
MH383468
MH383469
MH383470
0.80
MH383422
MH383423
MH383426
MH383460
C32_col
MH383425
MH383427
MH383428
MH383429
MH383431
MH383443
MH383447
MH383450
MH383451
MH383454
MH383455
MH383457
MH383458
MH383459
MH383461
0.70
0.80
C31_col
MH383435
MH383448
MH383464
0.86
MH383430
MH383449
MH383456
0.99